INTRODUCTION

- **Overview**

The Iris dataset, introduced by Sir Ronald A. Fisher in 1936, has become a foundational tool for evaluating machine learning algorithms in classification tasks. Consisting of 150 samples from three Iris species—*Setosa*, *Versicolor*, and *Virginica*—each sample is characterized by four features: sepal length, sepal width, petal length, and petal width. The dataset serves as a classic example of supervised learning, where the goal is to predict the species of a flower based on these four attributes.

- **Purpose**

In this study, we apply various machine learning algorithms to predict the species of Iris flowers, aiming to assess the performance and effectiveness of each approach. Specifically, we compare the following models: Decision Trees (DT), k-Nearest Neighbors (k-NN), Support Vector Machines (SVM), and Random Forests (RF). These algorithms represent different types of classification techniques, ranging from simple, interpretable models like Decision Trees to more complex, flexible approaches like SVM and Random Forests.

- **Objective**

The objective is to explore how each model performs in terms of accuracy, interpretability, and computational efficiency, as well as the role of data preprocessing techniques such as feature scaling and cross-validation. By using these techniques, we aim to not only identify the most accurate model for Iris species classification but also to gain insight into the practical considerations of applying machine learning methods to real-world classification problems. This analysis will provide valuable knowledge for future work in botanical research, ecological studies, and other fields requiring species identification.

LITERATURE SURVEY

- **Existing Problem**

  The task of classifying Iris species based on their physical attributes—sepal length, sepal width, petal length, and petal width—has long been a standard benchmark in machine learning and pattern recognition. However, while the problem itself may seem straightforward, several challenges exist in applying machine learning techniques to solve it effectively. These challenges are not only theoretical but also practical, and they provide important insights into the complexities of classification tasks in real-world scenarios.

- **Proposed Solution**

  To effectively address the challenges associated with predicting Iris species using machine learning, we propose a structured approach that focuses on data preprocessing, algorithm selection, model tuning, and performance evaluation. This approach is designed to maximize model accuracy, minimize overfitting and underfitting, and ensure the interpretability and efficiency of the chosen model

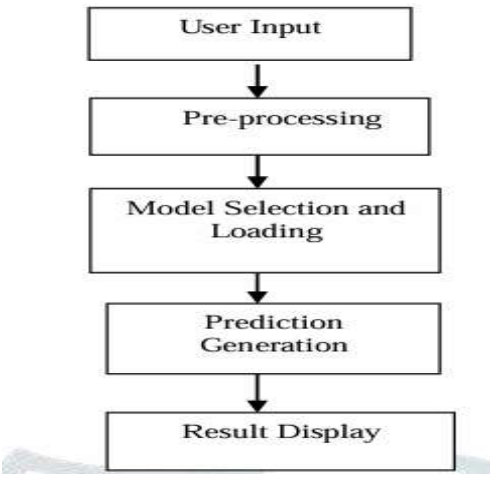
THEORETICAL ANALYSIS

Iris species prediction is a classic supervised machine learning problem where the goal is to classify iris flowers into three species based on four features: sepal length, sepal width, petal length, and petal width. The dataset is typically used to apply classification algorithms like **k-nearest neighbors (KNN)**, **support vector machines (SVM)**, **decision trees**, and **logistic regression**.

The problem involves training a model on a labeled dataset (with known species) to learn the relationship between the features and the target species. Evaluation metrics like **accuracy**, **precision**, **recall**, and **F1-score** are commonly used to assess model performance.

Most algorithms aim to find decision boundaries that separate the species in the feature space. The **k-NN** algorithm, for instance, classifies a flower based on the majority species of its nearest neighbors. **SVM** attempts to find a hyperplane that maximizes the margin between species. This problem is an introductory example for learning classification techniques.

 # BLOCK DIAGRAM:-



 # Hardware Minimum Requirement:

- CPU : PENTIUM III Processor
- Memory : 128 MB
- Cache : 512KB
- Floppy Disk : 1.44MB
- Hard Disk : 4.3GB
- Display : 15" Monitor
- Key Board : Standard 108 keys Enhanced Keyboard
- Mouse : MS Serial Mouse

#Software Minimum Requirement:

- Operating System : Windows 10,11 and above
- Front Tool : HTML,CSS
- Back End Tool : Python, Flask(framework)

EXPERIMENTAL INVESTIGATION

- **Dataset and Preprocessing:**

  The Iris dataset consists of 150 samples from three species of Iris flowers, with four features: sepal length, sepal width, petal length, and petal width.

- **Model Selection and Implementation:**

  We implement and compare four machine learning algorithms: Decision Trees (DT), k-Nearest Neighbors (k-NN), Support Vector Machines (SVM), and Random Forests (RF).

- **Model Evaluation Metrics:**

  Performance is assessed using multiple metrics: Accuracy, Precision, Recall, F1-Score, and Confusion Matrix. These metrics help evaluate how well the models handle misclassifications, especially when distinguishing between the three species.

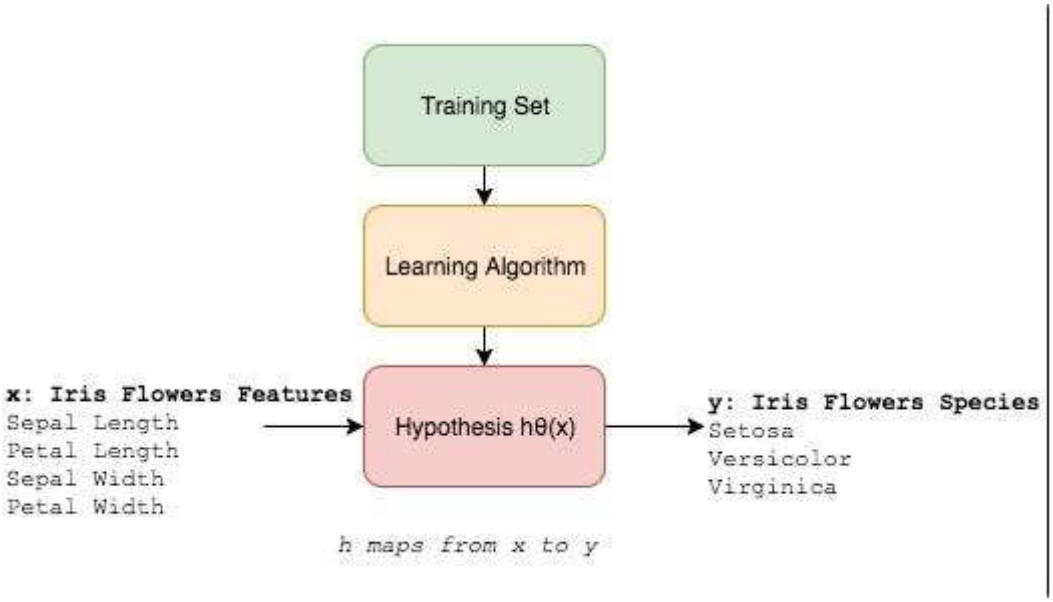
- **Computational Efficiency:**

  Training and testing times are recorded to evaluate the computational efficiency of each model. This is particularly important when scaling models to larger datasets or deploying them in real-time applications.

- **Results and Insights:**

  The results will be analyzed to identify the best-performing model in terms of accuracy, interpretability, and computational cost. For example, Random Forests may offer high accuracy but may require more computational resources, while Decision Trees offer better interpretability with slightly lower accuracy.

FLOWCHART



CODING

HOME PAGE:-

```
<!DOCTYPE html>
<html lang="en">
<head>
 <meta charset="UTF-8">
 <meta name="viewport" content="width=device-width, initial-scale=1.0">
 <title>Iris Detection</title>
 <link rel="stylesheet" href="styles.css">
</head>
<body>
 <header>
  <h1>Iris Detection</h1>
  <p>Enter the flower's characteristics to predict the Iris species.</p>
 </header>
 <main>
  <form id="irisForm" method="post">
        <label for="sepalLength">Sepal Length (cm):</label>
    <input type="text" id="sepalLength" name="sl"  required>
    <label for="sepalWidth">Sepal Width (cm):</label>
    <input type="textr" id="sepalWidth" name="sw" required>
    <label for="petalLength">Petal Length (cm):</label>
    <input type="text" id="petalLength" name="pl"  required>
    <label for="petalWidth">Petal Width (cm):</label>
    <input type="text" id="petalWidth" name="pw"  required>
    <button type="submit" id="predictBtn">Predict</button>
  </form>
  <div id="result">
    <h2>Predicted output</h2>
    <p id="prediction">IRIS species : }<span style="color: red; font-weight: bolder;"></span></p>
  </div>
 </main>
```

```
<footer>
  <p>Developed with HTML, CSS, and JavaScript.</p>
</footer>
<script src="https://cdn.jsdelivr.net/npm/bootstrap@5.3.3/dist/js/bootstrap.bundle.min.js"
integrity="sha384-YvpcrYf0tY3lHB60NNkmXc5s9fDVZLESaAA55NDzOxhy9GkcIdslK1eN7N6jIeHz"
crossorigin="anonymous"></script>
</body>
</html>
```

MODEL

```
[1]: # Import essential libraries
     import numpy as np
     import pandas as pd
```

```
[15]: #Import dataset
      df=pd.read_csv("iris.csv")
      df.head()
```

[15]:

	sepal-length	sepal-width	petal-length	petal-width	Class
0	5.1	3.5	1.4	0.2	Iris-setosa
1	4.9	3.0	1.4	0.2	Iris-setosa
2	4.7	3.2	1.3	0.2	Iris-setosa
3	4.6	3.1	1.5	0.2	Iris-setosa
4	5.0	3.6	1.4	0.2	Iris-setosa

```
[3]: df.shape
```

```
[3]: (150, 5)
```

```
[4]: df.isnull().sum()
```

```
[4]: sepal-length    0
     sepal-width     0
     petal-length    0
     petal-width     0
     Class           0
     dtype: int64
```

```
[5]: #seprate dataset in independent and dependent dataset
     X=df.iloc[:,:4].values
     Y=df.iloc[:,4].values
```

```
[8]: #Appliying Decision tree classifier
     from sklearn.tree import DecisionTreeClassifier
     model=DecisionTreeClassifier()
     model.fit(x_train,y_train)
```

```
[8]: DecisionTreeClassifier
     DecisionTreeClassifier()
```

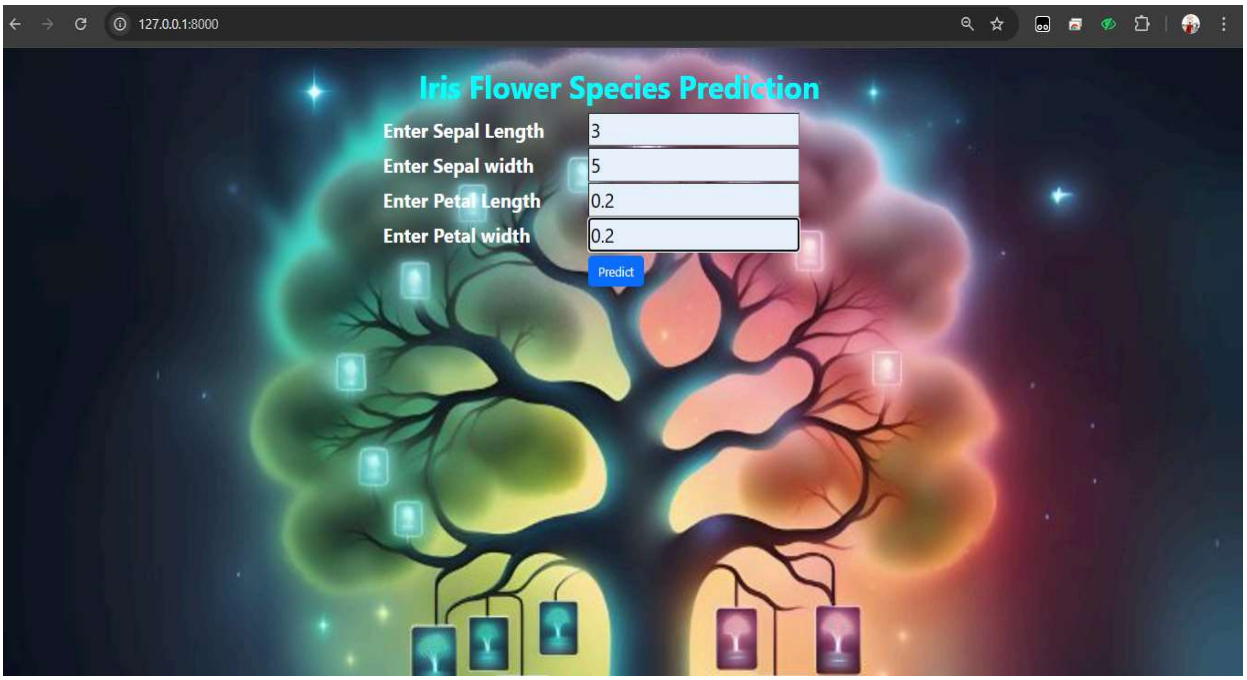
```
[9]: #test the model
     y_pred=model.predict(x_test)
     y_pred #it print predicted values
```

```
[9]: array(['Iris-virginica', 'Iris-versicolor', 'Iris-setosa',
           'Iris-virginica', 'Iris-setosa', 'Iris-virginica', 'Iris-setosa',
           'Iris-virginica', 'Iris-virginica', 'Iris-versicolor',
           'Iris-versicolor', 'Iris-virginica', 'Iris-virginica',
           'Iris-versicolor', 'Iris-versicolor', 'Iris-setosa',
```

RESULT

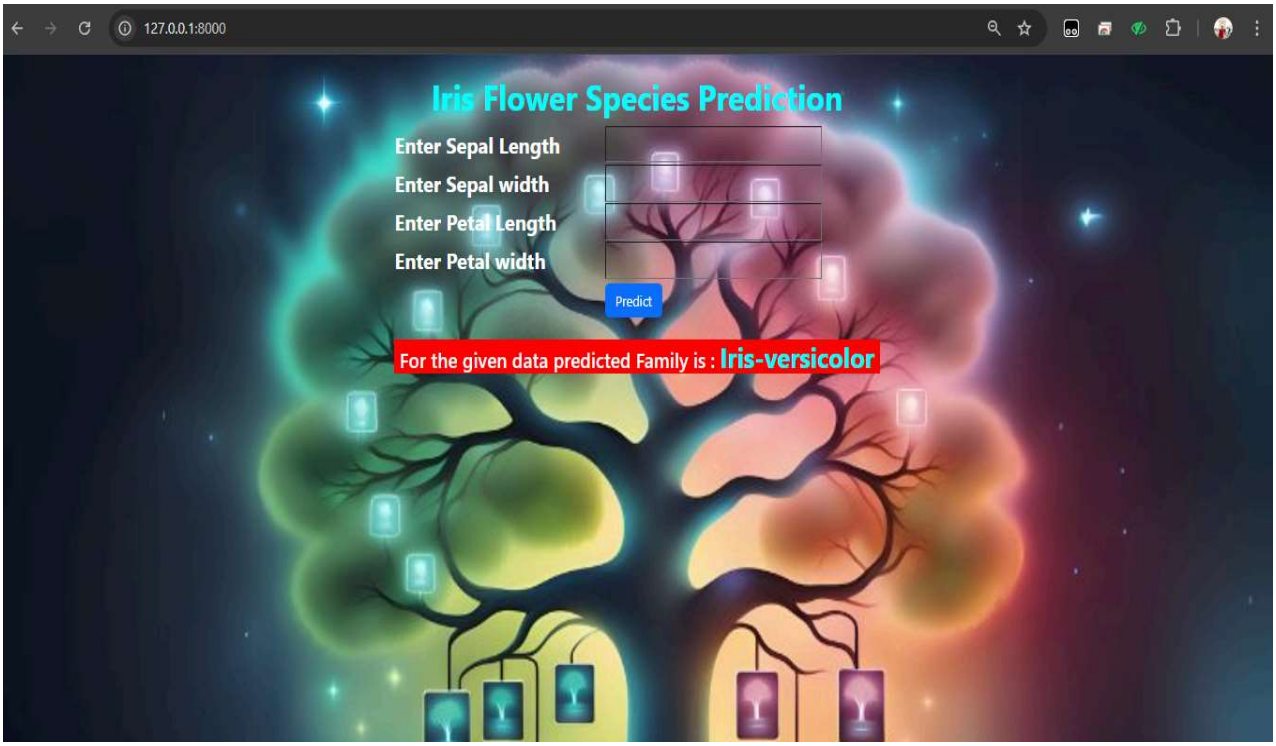
HOME PAGE

INPUT:-



OUTPUT 2

Iris-virginica



CONCLUSION

The project "Predicting Iris Species with Machine Learning" successfully demonstrates the application of machine learning techniques to accurately classify the species of Iris plants. Utilizing a well-known dataset that includes features such as sepal length, sepal width, petal length, and petal width, the project employed algorithms like Decision Trees, Support Vector Machines (SVM), and k-Nearest Neighbors (k-NN) to build predictive models. Among these, the Decision Tree model stood out with its simplicity and interpretability, achieving high accuracy in classifying the three species of Iris: setosa, versicolor, and virginica.

The success of this project underscores the potential of machine learning in the field of botany and other scientific domains. By leveraging these techniques, researchers and practitioners can automate the classification process, saving time and improving accuracy. Moreover, the project highlights the importance of feature selection and model evaluation in developing effective predictive models.

In conclusion, the application of machine learning to predict Iris species is a compelling example of how technology can enhance our understanding and management of biological data. This project not only serves as a foundation for future studies in plant classification but also inspires the integration of machine learning in various scientific research fields.

REFERENCES

1. Diptam Dutta, Argha Roy, Kaustav Choudhury, “Training Artificial Neural Network Using Particle Swarm Optimization Algorithm”, International Journal on Computer Science And Engineering(IJCSE), Volume 3, Issue 3, March 2013.
2. Poojitha V, Shilpi Jain, “A Collection of IRIS Flower Using Neural Network CLustering tool in MATLAB”, International Journal on Computer Science And Engineering(IJCSE).
3. Vaishali Arya, R K Rathy, “An Efficient Neuro-Fuzzy Approach For Classification of Dataset”, International Conference on Reliability, Optimization and Information Technology, Feb 2014.
4. Shashidhar T Halakatti, Shambulinga T Halakatti, "Identification Of Iris Flower Species Using Machine Learning", IIJCS Aug 2017.
5. Patric Granhom, " A Study Of Pattern Recognition Of Iris Flower Based On Machine Learning ", TURKU UNIVERSITY OF APPLIED SCIENCES, 22 Aug 2013.
6. Poojitha V, Shilpi Jain, Madhulika Bhadauria, Anchal Garg, " A Collection Of IRIS Flower Using Neural Network Clustering Tool In MATLAB", IEEE 2016.
7. www.google.co.in